GB2318472

Publication Title:

Generating inter-frame encoded video signals

Abstract:

Abstract of GB2318472

A frame of time-domain video signals Si is supplied via a frame reorderer 40 to a motion vector estimator 41 and then to a DCT transformer 43 and a quantiser 44 to produce a quantised, transformed current frame S T iq. A subtracter 42 subtracts a transformed quantised prediction frame S T pq from S T iq to produce an inter-encoded signal * capital Greek delta *. The prediction frame S T pq is produced by adding the prediction frame S T pq to the inter-encoded signal * capital Greek delta *, dequantising 46 and inverse transforming 47 the sum and applying it to a motion compensation circuit 49 together with the output of the motion vector estimator 41 to produce a time domain prediction signal Sp. This is then transformed 50 and quantised 51 to produce S T pq. The quantisers 44 and 51 are controlled by the same quantisation level Qv.

(12) UK Patent Application (19) GB (11) 2 318 472 (13) A

(43) Date of A Publication 22.04.1998

(21) Application No 9715918.4

(22) Date of Filing 28.07.1997

(30) Priority Data
(31) 9621067 (32) 09.10.1996 (33) GB

(71) Applicant(s)
Sony United Kingdom Limited

(Incorporated in the United Kingdom)

The Heights, Brooklands, WEYBRIDGE, Surrey, KT13 0XW, United Kingdom

(72) Inventor(s)
James Hedley Wilkinson

(74) Agent and/or Address for Service
D Young & Co
21 New Fetter Lane, LONDON, EC4A 1DA, United Kingdom

(51) INT CL[6]
H04N 7/50

(52) UK CL (Edition P)
H4F FD12X FD3D FD3P FD3T FD30B FD30H FD30K FD30R FRG

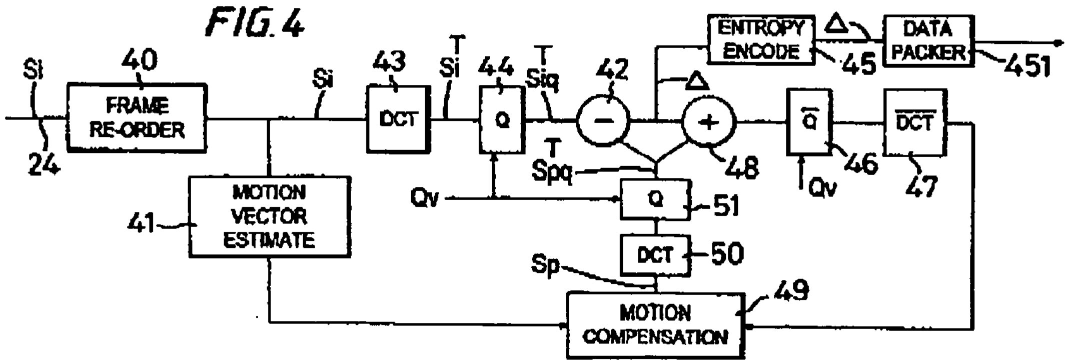
(56) Documents Cited
EP 0540358 A2 EP 0535960 A2 US 5440344 A

(58) Field of Search
UK CL (Edition O) **H4F FRG**
INT CL[6] **H04N 5/926 7/46 7/50**
Online: **WPI, INSPEC**

(54) **Generating inter-frame encoded video signals**

(57) A frame of time-domain video signals Si is supplied via a frame reorderer 40 to a motion vector estimator 41 and then to a DCT transformer 43 and a quantiser 44 to produce a quantised, transformed current frame S^Tiq. A subtracter 42 subtracts a transformed quantised prediction frame S^Tpq from S^Tiq to produce an inter-encoded signal Δ. The prediction frame S^Tpq is produced by adding the prediction frame S^Tpq to the inter-encoded signal Δ, dequantising 46 and inverse transforming 47 the sum and applying it to a motion compensation circuit 49 together with the output of the motion vector estimator 41 to produce a time domain prediction signal Sp. This is then transformed 50 and quantised 51 to produce S^Tpq. The quantisers 44 and 51 are controlled by the same quantisation level Qv.

FIG. 4



GB 2 318 472 A

At least one drawing originally filed was informal and the print reproduced here is taken from a later filed formal copy.

1/7

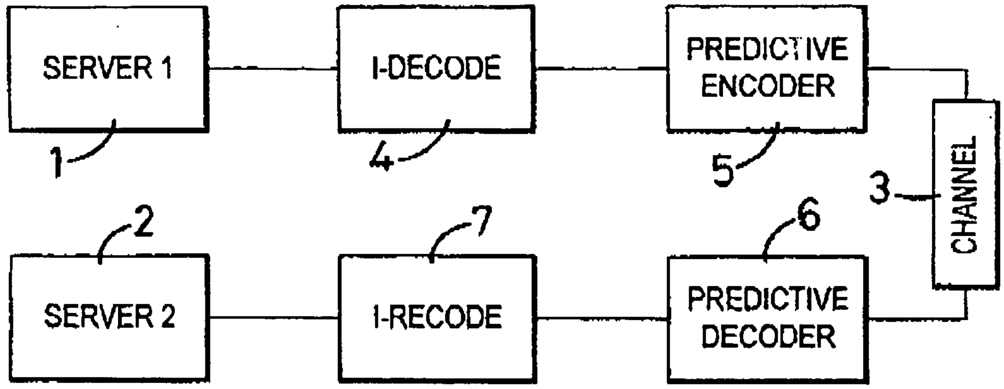


FIG.1

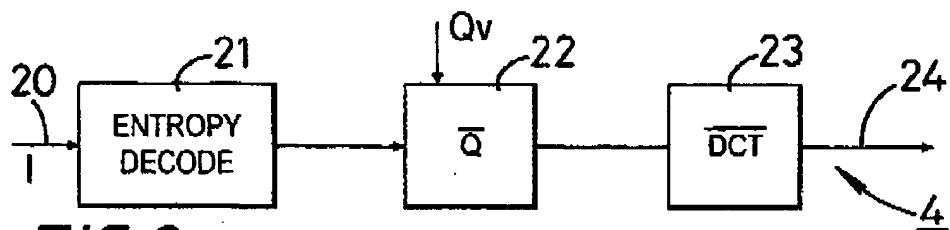


FIG.2

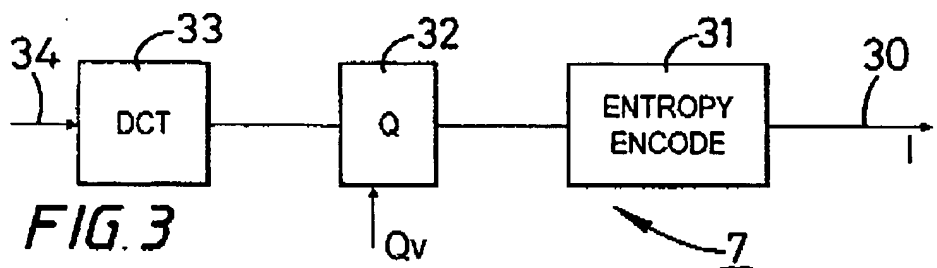


FIG.3

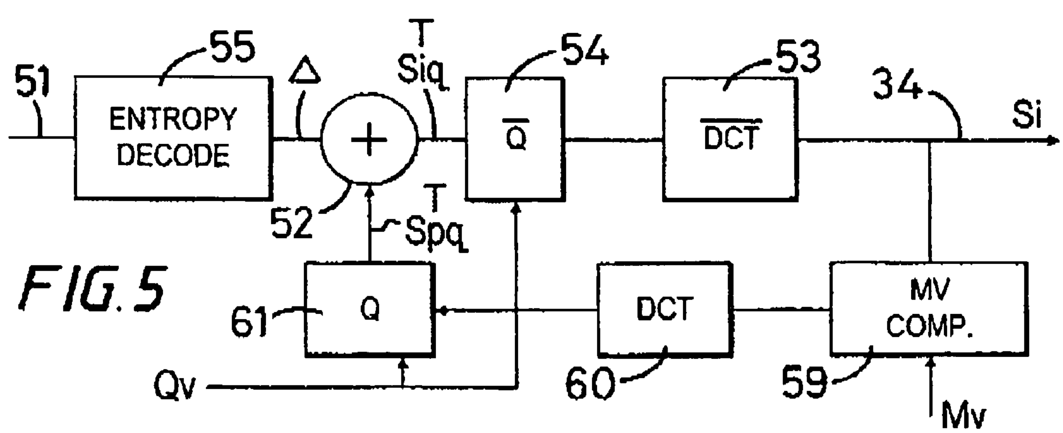
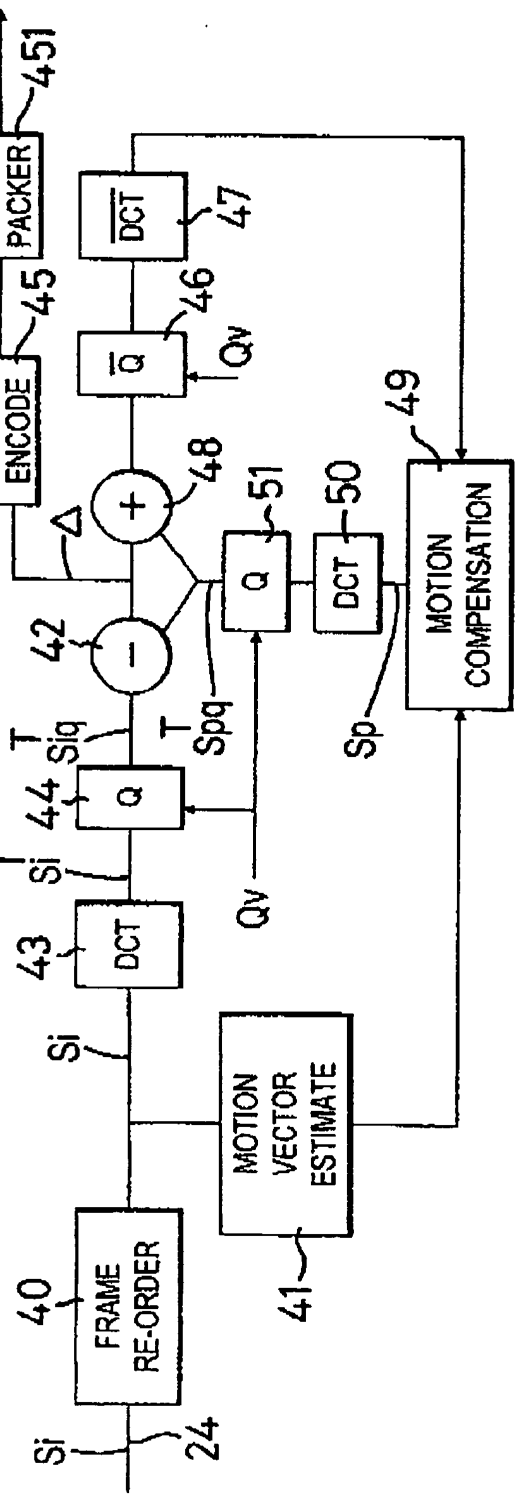


FIG.5

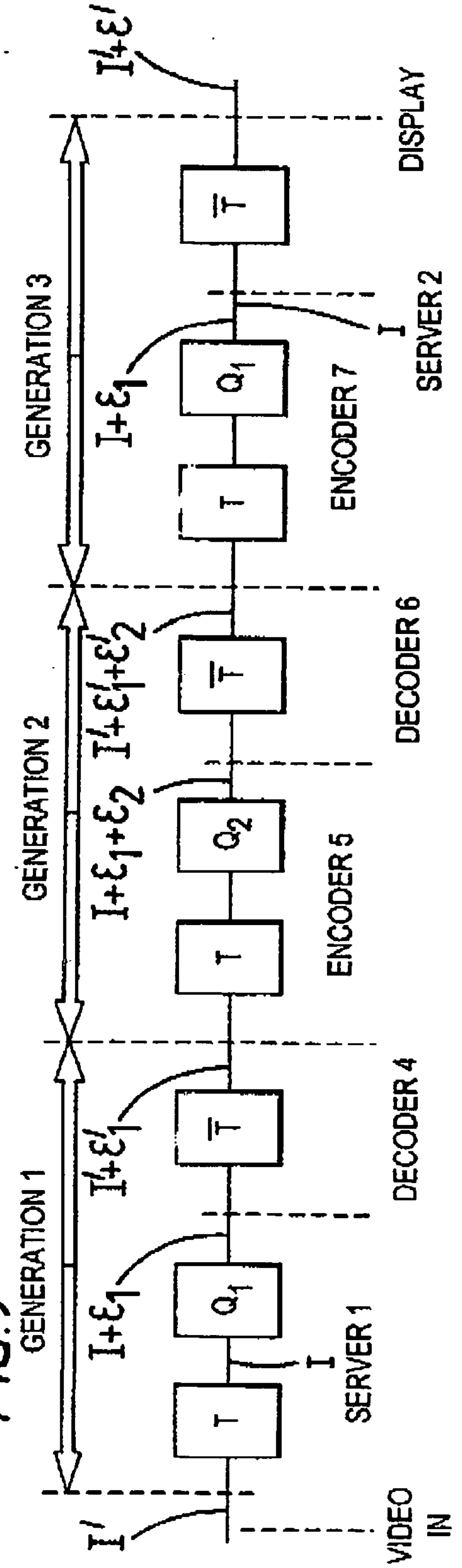
2/7

FIG. 4



Si
24
FRAME RE-ORDER
40
MOTION VECTOR ESTIMATE
41
Si
DCT
43
Si
Q
44
Siq
Qv
Spq
42
Sp
DCT
50
Q
51
MOTION COMPENSATION
49
+
48
Δ
ENTROPY ENCODE
45
Δ
Q̄
46
Qv
DCT̄
47
DATA PACKER
451

FIG. 9



GENERATION 1
GENERATION 2
GENERATION 3
I'
VIDEO IN
T
I
SERVER 1
Q1
I+ε1
I'+ε'1
T̄
DECODER 4
T
Q2
I+ε1+ε2
ENCODER 5
I'+ε'1+ε'2
T̄
DECODER 6
T
Q1
I+ε1
ENCODER 7
I
SERVER 2
T̄
I'+ε'
DISPLAY

3/7

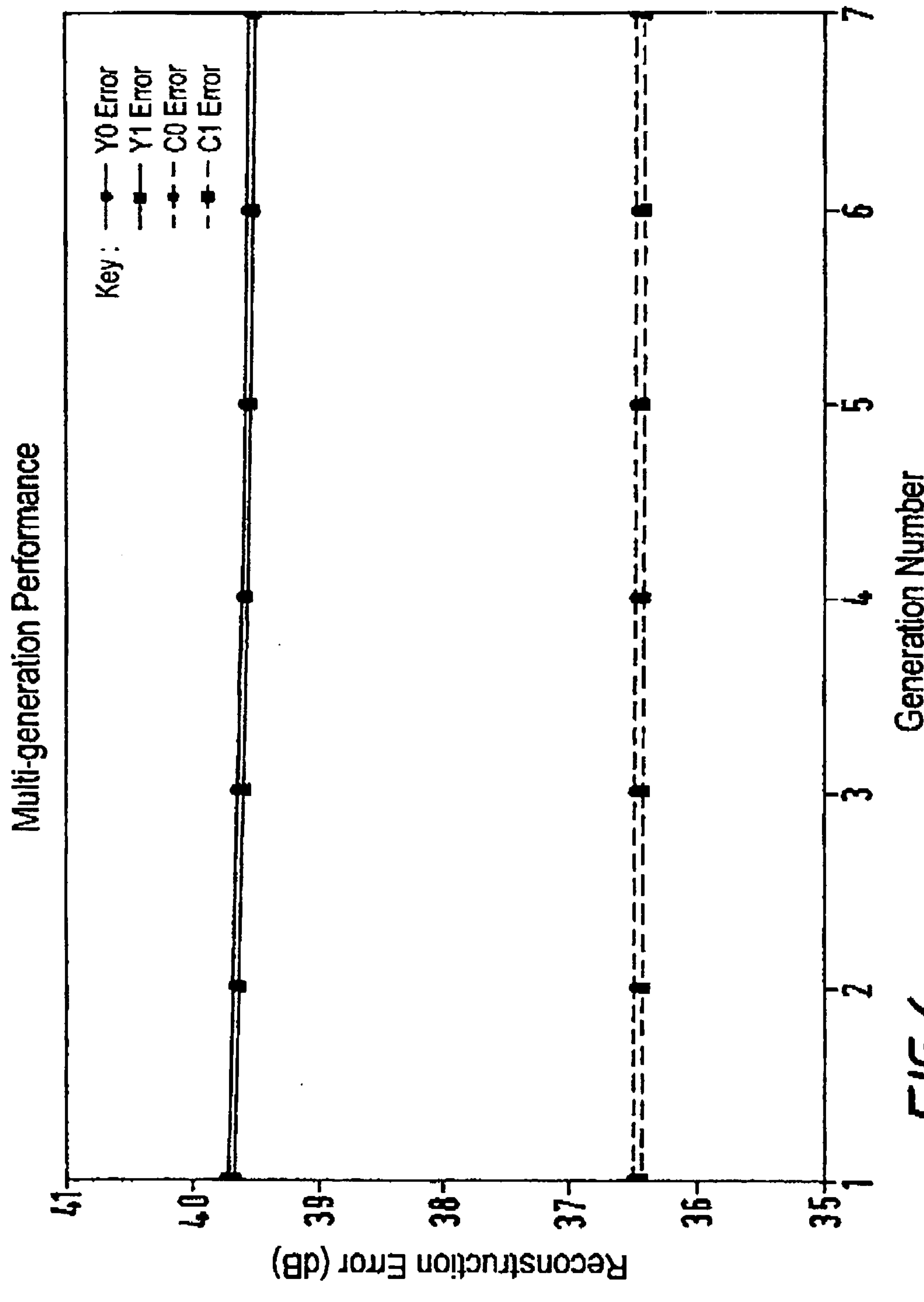
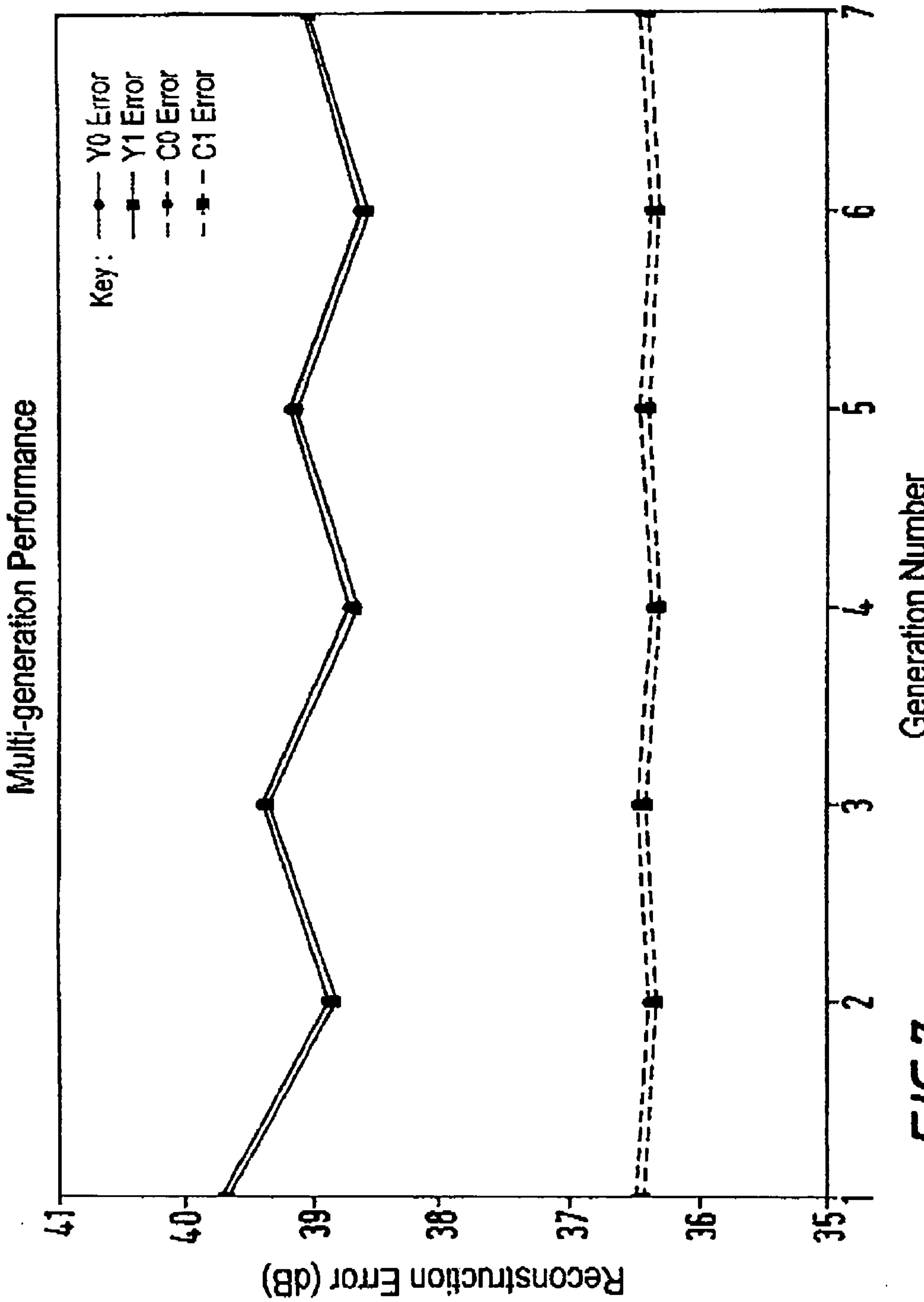

Multi-generation Performance
Key :
Y0 Error
Y1 Error
C0 Error
C1 Error
Reconstruction Error (dB)
41
40
39
38
37
36
35
Generation Number
1
2
3
4
5
6
7

FIG.6



Multi-generation Performance
Key:
Y0 Error
Y1 Error
C0 Error
C1 Error
Reconstruction Error (dB)
41
40
39
38
37
36
35
Generation Number
1
2
3
4
5
6
7

FIG.7

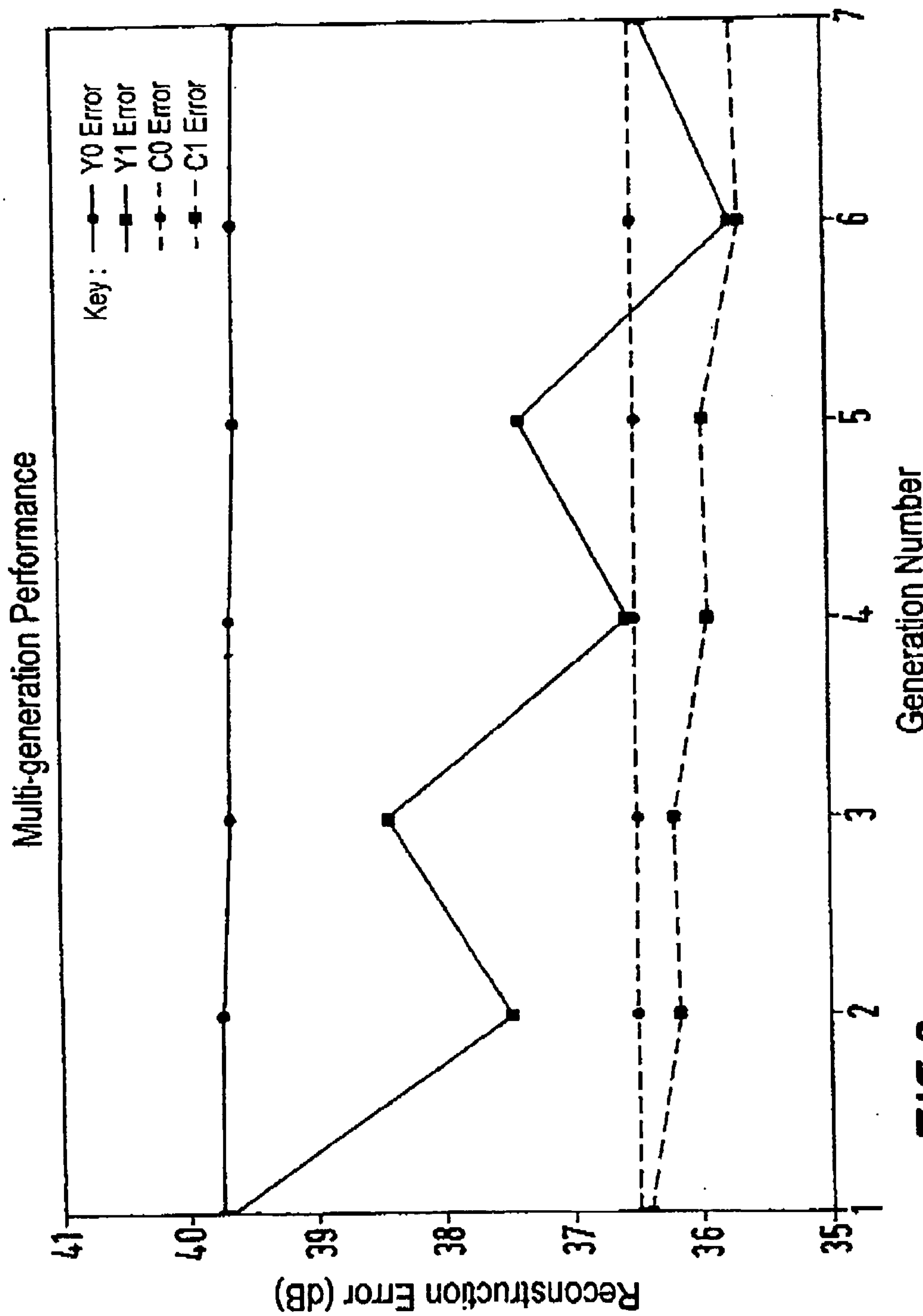

Multi-generation Performance
Key :
Y0 Error
Y1 Error
C0 Error
C1 Error
Reconstruction Error (dB)
41
40
39
38
37
36
35
1
2
3
4
5
6
7
Generation Number

FIG.8

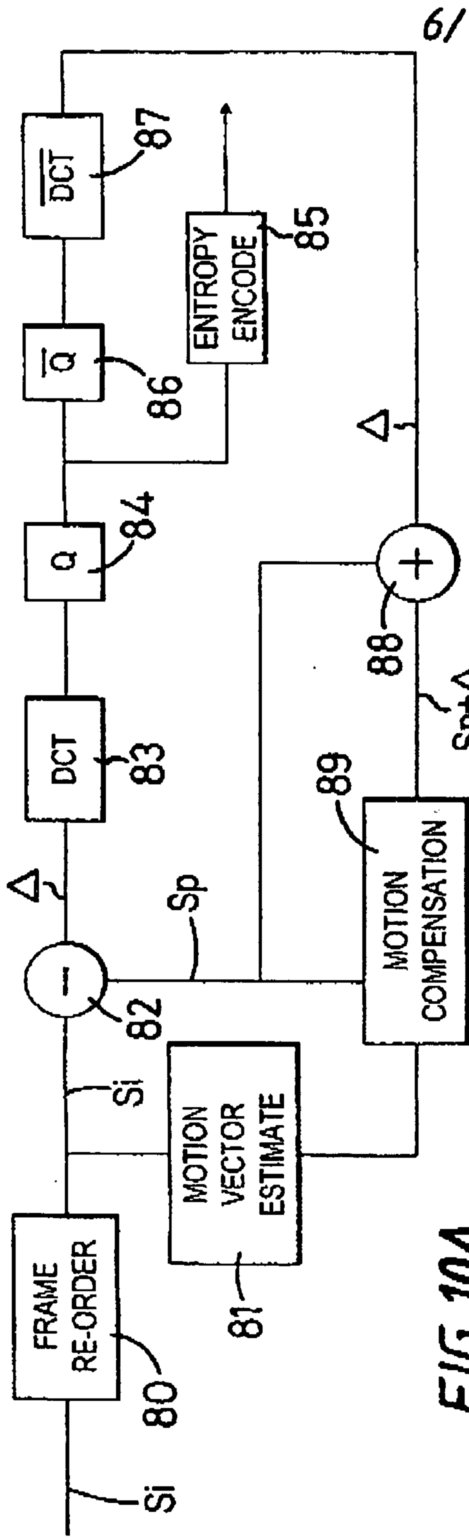


FIG. 10A

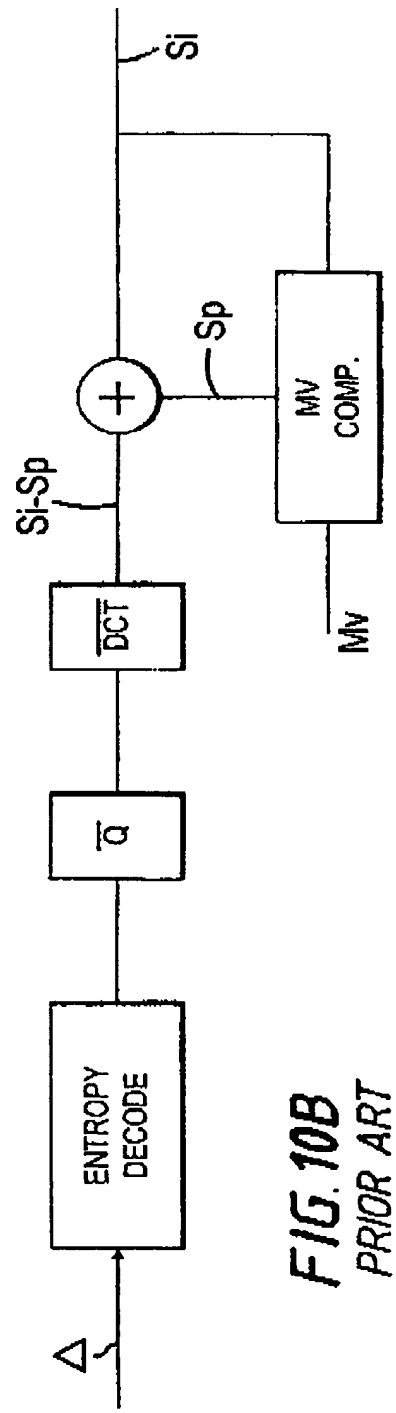


FIG. 10B
PRIOR ART

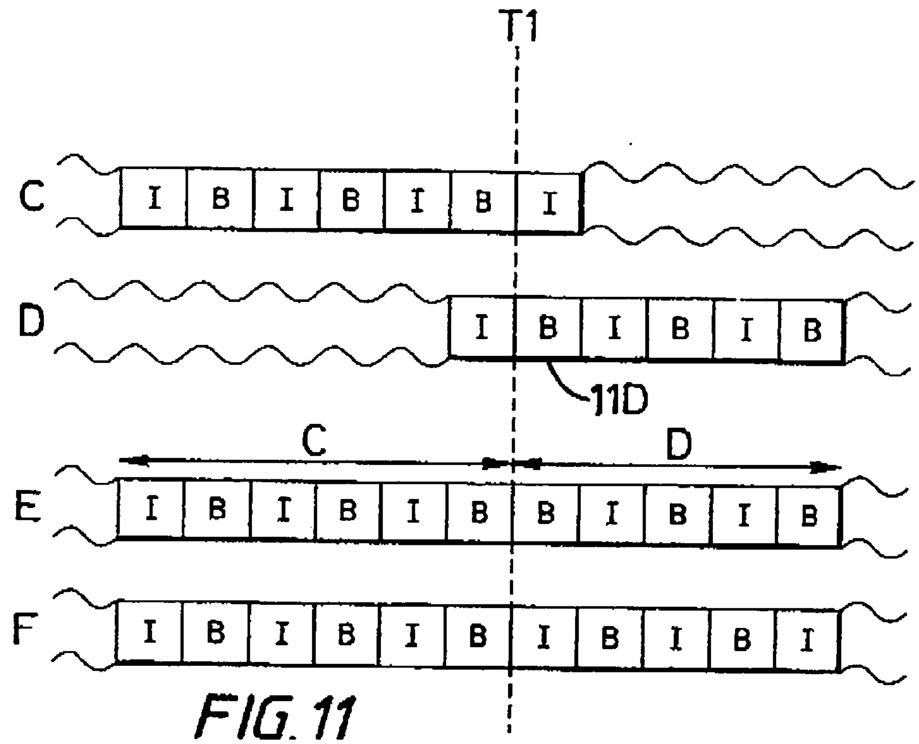

T1
C
I B I B I B I
D
I B I B I B
11D
C
D
E
I B I B I B B I B I B
F
I B I B I B I B I B I

FIG. 11

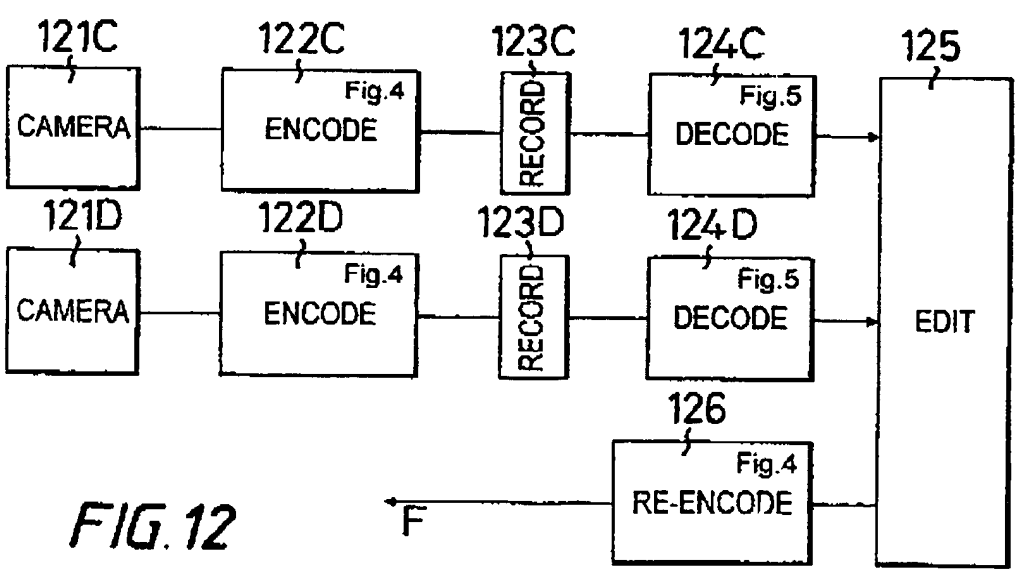

121C
CAMERA
122C
Fig.4
ENCODE
123C
RECORD
124C
Fig.5
DECODE
125
EDIT
121D
CAMERA
122D
Fig.4
ENCODE
123D
RECORD
124D
Fig.5
DECODE
126
Fig.4
RE-ENCODE
F

FIG. 12

2318472

PROCESSING ENCODED SIGNALS

The present invention relates to the processing of encoded signals.

It is known to digitally encode signals according to the MPEG 2 standard defined in ISO/IEC 13818-2. A signal encoded according to MPEG 2 comprises I-frames and P and/or B-frames. I-frames or intra-encoded frames are frames in which all the information necessary to decode the frame is preserved within the frame. P and B-frames are inter-encoded frames. P and B-frames contain motion-compensated differences between two-frames. In practice the difference is between a frame and a prediction derived from one or two other frames. The differences are ultimately referenced to one or two I-frames. P-frames are predicted from a single preceding reference frame which may be an I-frame or a P-frame. B-frames are bidirectionally interpolated from the nearest preceding I or P frame and the nearest succeeding I or P frame. The encoded frames are organised into a Group of Pictures or GOP of which at least one frame is an I-frame, and which may include one or more B and/or P frames.

Figure 10A is a block diagram of a typical MPEG 2 signal encoder. The operation of such an encoder is well known and therefore will be described only briefly herein.

The compressor of Figure 10A comprises a frame re-orderer 80, a motion vector estimator 81, a motion compensator 89, a subtracter 82, an adder 88, a DCT transform unit 85, a quantiser 84, an entropy encoder 85, an inverse quantiser 86 and an inverse DCT unit 87.

As is known in MPEG, blocks of transformed data are organised into macroblocks and the macroblocks are assembled into slices and slices into pictures. There are the three types of picture, I, B and P described above.

The frame reorderer 80 receives input video data and acts on successive groups of pictures (GOP) to reorder the pictures so that each picture within the GOP is compressed *after* those pictures on which it depends. For example, if a B-picture (bi-directionally predicted picture) depends on a following I- or P- picture, it is reordered to be compressed after that I- or P- picture.

For example, if a GOP comprises the following four initial frames (in the order in which they are *displayed*), $I_0B_1B_2P_3\ldots$, where the P-picture uses the I-picture as a reference and the two B- pictures use the surrounding I- and P-pictures as references, then the frame reorderer 80 will reorder the GOP to be *compressed* in the following order: $I_0P_3B_1B_2\ldots$ The pictures are also transmitted to the decoder for decoding in that order $I_0P_3B_1B_2\ldots$

I- pictures are intra-picture encoded, that is to say the encoding is not based on any other reference pictures. An I- picture in a GOP is therefore passed from the frame reorderer 80 to the DCT transform unit 85, the quantiser 84 and the entropy encoder 85 to generate output compressed data representing that I- picture.

The compressed I-picture data is also passed from the quantiser 84 through a decompression chain formed by the inverse quantiser 86, and the inverse DCT transform unit 87. This reconstructs a version of the I- picture present in the decoder which is passed to the motion predictor 540.

The next picture of the GOP to be compressed, which will generally be a P-picture which depends on the I- picture as a reference, is passed from the frame reorderer 80 to the motion vector estimator 81 which generates motion vectors indicative of image motion between the I- and P- pictures. The motion compensator 89 then generates a predicted version of the P picture using the motion vectors and the decoded version of the I- picture. This predicted version of the P- picture is subtracted from the actual P- picture by the subtracter 82 and the difference between the two frames is passed to the DCT transform unit 85 and the quantiser 84 and the coder 85 for compression. As before, the encoded (compressed) difference data is output by the quantiser 85 and is then decoded by the decompression chain 86, 87 to regenerate a version of the difference data.

In the adder 88 the difference data is then added to the previously decompressed version of the I- picture to generate a decompressed version of the P-picture which is then stored in the motion predictor 89 for use in the compression of the next picture.

This process continues, so that each picture which uses other pictures as a reference is in fact compressed by encoding difference data between the input picture

and a predicted version of the input picture formed by motion prediction from a previously compressed and then decompressed version of the reference picture. By using a predicted picture derived by decompressing and dequantizing, reconstruction errors are reduced in the finally displayed image. This is because the difference data is then the "error" between the input picture and the picture data which will be available at the decompressor to be added to the error to reconstruct the input picture.

Information identifying the type of picture I, B, P and information relating to quantisation levels and motion vectors is conveyed as syntax data in headers associated with e.g. the macro blocks.

The corresponding decoder is shown in Figure 10B from which it is apparent the predicted picture Sp is added to the decoded difference data Δ in the time domain to produce the decoded picture Si.

If the same video signal is MPEG encoded, and decoded more than once, and if the data rate (quantisation level) changes between generations errors in the reconstructed signal (i.e. the differences between the original video signal and the signal produced following one or more generations of encoding and decoding) increases with each generation.

Such encoding decoding and re-encoding and decoding over several generations may occur in for example editing systems using different encoded data rates at each generation. Such encoded signals are decoded to the time domain for frame accurate editing and then re-encoded. Such repeated, decoding and re-encoding may occur also in other circumstances.

British Patent Application 9621067.9 filed 9 October 1996 (Attorney reference: I-96-1; P/749.GB; S96P5032GB00) proposes a method of processing digitally encoded signals comprising intra-encoded frames and inter-encoded frames, the inter-encoded frames having associated therewith respective quantisation levels and respective motion vectors, the method comprising converting the inter-encoded frames to intra-encoded frames retaining unchanged the respective quantisation levels. It has been found that, by retaining, unchanged, the respective quantisation levels, the reconstruction error of intra-encoded frames is very small over many generations of decoding and re-encoding. This is shown in accompanying Figure 8 where Yo and

Co indicate the luminance and chrominance reconstruction errors for I-frames. Although there is some increase in error over many generations it is small. The figure is based on using codec 'A' which encodes and decodes I-B 2-frame GOP followed by a codec 'B' which encodes and decodes a single frame GOP (I-frame only). The results in Figure 8 show the effects of cascading through a succession of A, B codecs thus: A, B, A, B, A etc. In Figure 8, lines Y_1 and C_1 indicate error for B frames. As shown by Y_1 and C_1 the error increases using the B codec but then recovers to close to the first generation performance when followed by a third generation of codec 'A'.

Figure 8 indicates the reconstruction error as signal to noise ratio (dB), decreasing signal to noise ratio indicating increasing reconstruction error.

The present invention seeks to provide signal encoder said decoders which allow a reduction in reconstruction error when signals are repeatedly encoded and decoded.

According to one aspect of the present invention there is provided a method of and apparatus for encoding video signals in which inter-encoded pictures and intra-encoded pictures are produced from an input video signal comprising a sequence of pictures, wherein each inter-encoded picture is produced by forming the difference between:

a) a transformed and quantised picture of the input signal; and

b) a corresponding separately transformed and quantised predicted pictured derived from the input signal;

the quantisations applied to the transformed picture of the input signal and to the transformed predicted picture being the same.

According to another aspect of the present invention, there is provided a method of and apparatus for decoding video signals encoded according to the method of said one aspect, in which each inter-encoded picture is decoded by forming the sum of:

a) the transformed and quantised inter-encoded picture; and

b) a separately transformed and quantised predicted picture; and

dequantising and inverse transforming the sum to produce the decoded picture;

the quantisation applied to quantise the transformed predicted picture and to dequantise the said sum being similar to or the same as the quantisation of the inter-encoded picture.

According to a further aspect of the invention there is provided a method of or apparatus for processing input video signals comprising: encoding the signals according to the method of said one aspect, to produce encoded signals; applying the encoded signals to a transmission channel, storage device or utilization means and receiving the encoded signals therefrom; and decoding the said received encoded signals according to the method of said another aspect to produce decoded picture signals.

The video signals are encoded so that their bit rate matches that of the transmission channel, storage device or utilization means, and then decoded when received therefrom. The quantisation level may be changed during encoding to achieve the desired bit rate or may be held constant. The above-mentioned aspects of the invention provide reduced reconstruction error by forming inter-encoded pictures from separately transformed and quantised pictures which are effectively I-pictures and are in the transform domain, and subject to the same quantisation.

According to yet another aspect of the invention, there is provided a method of or apparatus for processing a method of processing video signals in which at least some intra-encoded pictures are converted to inter-encoded pictures and then back to intra-encoded pictures, the quantization levels of the pictures which are converted and reconverted being maintained constant.

The yet another aspect of the invention provides reduced reconstruction error by maintaining the quantisation level constant as I pictures are converted to P and/or B pictures and reconverted back to I.

For a better understanding of the present invention reference will now be made, by way of example, to the accompanying drawings, in which:

Figure 1 is a schematic block diagram of a system, incorporating an embodiment of the invention, for transferring encoded image data from a first video server to a second video server via a channel;

Figure 2 is a schematic block diagram of an Intra-frame decoder, useful in the

system of Figure 1;

Figure 3 is a schematic block diagram of an Intra-frame encoder useful in the system of Figure 1;

Figure 4 is a schematic block diagram of an illustrative predictive encoder according to the present invention and useful in the system of Figure 1;

Figure 5 is a schematic block diagram of an illustrative predictive decoder according to the present invention and useful in the system of Figure 1;

Figure 6 is a graphical representation of reconstruction error using an encoder/decoder according to the present invention under a first set of conditions on an image sequence known as "flower garden";

Figure 7 is a graphical representation of reconstruction error using an encoder/decoder according to the present invention under a second set of conditions on the image sequence known as "flower garden";

Figure 8 is a graphical representation of the reconstruction error for MPEG encoding and decoding of the image sequence known as "flower garden"; and

Figure 9 is a diagram useful in explaining the results of Figure 6, 7 and 8.

Figure 10A is a schematic block diagram of a known MPEG 2 signal encoder;

Figure 10B is a schematic block diagram of a corresponding decoder;

Figure 11 is a schematic diagram of waveforms useful in explaining a futher application of the present invention; and

Figure 12 is a block diagram of the further application.

Referring to Figure 1, an illustrative system for transferring digitally encoded image data from a first video server 1 to a video second server 2 via a channel 3 is shown. The channel 3 may be a signal transmission system and/or include a signal storage device such as a video tape player/reproducer and/or be some other utilization means such as a computer.

In this example video servers 1 and 2 store the video encoded as I-frames only. That allows, for example, frame accurate editing of the encoded video.

The channel 3 has a bandwidth, or data rate, too small to allow the I-frames to be transferred from server 1 to server 2 at the normal video frame rate.

The system further comprises an I-decoder 4 which decodes the I-frames to

the time domain or digital baseband, a predictive encoder 5 according to the invention which re-encodes the frames at a bit-rate which matches that of the channel 3, a corresponding predictive decoder which again decodes the frames to the time domain 6 and an I-encoder 7 which reconstructs the original I-frames for storage on the second server 2.

An example of an I-decoder 4 is shown in Figure 2. It comprises an input 20 for received I-frames from the first server 1, an entropy decoder 1, a dequantiser 22, an inverse DCT transformer 23 and an output 24 at which a digital baseband video signal is produced.

The corresponding encoder 7 of Figure 3 comprises an input 34 for receiving a digital baseband signal, a DCT transformer 33, a quantiser 32, an entropy encoder 31 and an output 30 at which I-frames are produced.

The operation of the decoder 4 and encoder 7 are conventional and will not be described here.

Figure 4 shows an illustrative predictive encoder according to the invention. Although the encoder does not comply with MPEG 2 much of its operation is the same or similar to that of an MPEG 2 encoder and it produces encoded signals capable of being decoded by an MPEG decoder. The encoder comprises an input 24 for receiving the decoded I-frames from the decoder 4, a frame reorderer 40 which reorders the frames, a motion vector estimator 41, a DCT transformer 43, a quantiser 44, a subtracter 42 in which the difference $\Delta = (S^Tiq - S^Tpq)$ between a current transformed and quantised frame S^Tiq and a predicted transformed and quantized frame S^Tpq is formed, and an entropy encoder 45 which entropy encodes the difference Δ to produce the desired encoded output signal. The predicted frame S^Tpq is produced by a decompression chain including an adder 48 which adds the predicted frame S^Tpq to the difference frame Δ, a dequantiser 46, and an inverse DCT transformer 47, and by a motion vector compensator 49, a DCT transformer 50 and a quantiser 51. In accordance with this example of the present invention, the quantisers 44 and 51 quantise to the same values Qv.

It will be noted that unlike the MPEG encoder of Figure 8 which forms the difference frame Δ in the time domain, the encoder of Figure 4 forms the difference

frame ▵ in the transform domain. It will also be noted that unlike the MPEG encoder of Figure 8, in which the dequantisation and the inverse DCT transformer operate on the transformed difference frame ▵, in the encoder of Figure 4 the dequantiser 46 and inverse DCT transformer 47 operate on the transformed current frame S^Tiq.

In operation, the decoded signal Si received at input 24 is reordered into a GOP comprising an I-frame and one or more P and/or B frames. The quantisation values Qv applied to the frames, the GOP length, and the number of P and B frames in the GOP is chosen to achieve a bit rate which matches that of the channel 3. The bit rate may be selected to be less than that of the channel and additional bits may be inserted into the bit stream using the known technique of bit justification so that the bit rate matches that of the channel 3. The bit justification occurs in a conventional data packer 451 after the entropy encoder 45.

The quantization values Qv and the motion vectors Mv are included in syntax data included in the encoded bitstream in the same way as in MPEG 2 encoded signals.

The corresponding decoder of Figure 5 comprises an input 51 for receiving the encoded signal ▵ from encoder 5 via the channel 3, an entropy decoder 55, an adder 52, a dequantiser 54, an inverse DCT transformer 43, a motion vector compensator 59, a DCT transformer 60, a quantiser 61 and an output 34, at which the decoded signal Si is produced.

In operation, the input signal ▵ is entropy decoded and added in adder 52 to a transformed and quantized prediction signal S^Tpq to produce a transformed and quantized signal S^Tiq which is dequantised in dequantiser 54 and inverse transformed in block 53 to produce the time domain decoded signal Si which is the desired output signal. The prediction signal S^Tpq is produced by applying motion vectors M_v to the decoded signal Si in a motion compensator 59, and DCT transforming 60 and quantising 61 the result. The quantisation values Qv applied to quantiser 61 and to the dequantizer 54 are the same, and are the values Qv applied to the transformed difference signal ▵ in the encoder and transmitted to the decoder as syntax data.

It will be noted that unlike the MPEG decoder of Figure 10B in which the prediction signal Sp is a time domain signal and is added to the dequantised time

domain difference signal Δ, in the decoder of Figure 5 both the difference signal Δ and the prediction signal S^Tp are in the transform domain and are added together in the transform domain. It is this aspect which is the key component to ensuring good cascade performance because the signals in elements 43, 44, 45 and 47 are intra-frame coded.

Thus in the decoder of Figure 5, the dequantisation 54 operates on the transformed intra-encoded signal S^Ti, not the transformed difference signal Δ as in the MPEG decoder of Figure 8B.

The system of Figure 1 using the encoder and decoder of Figures 4 and 5 may operate in two modes. In a first one of the modes the quantisation levels of the original I frames are maintained constant throughout the encoding and decoding process. In that case as illustrated by Figure 6, lines Yo and Co representing frames originally encoded as I-frames, suffer very little reconstruction error (−0dB) when decoded to the time domain and re-encoded to I over several generations, where one generation is for example one cycle of decode I to time domain and re-encode back to I. Lines Y_1 and C_1 show for the same mode that B-frames may be similarly decoded and re-encoded over several generations with substantially the same reconstruction error as I-frames.

In a second of the two modes, the quantisation level of the signal produced by the encoder 5 is changed (as compared to that of the original I-frames supplied by the server 1) for example to produce a bit rate which matches that of the channel.

Figure 7 shows that when I and B frames are decoded and re-encoded over several generations, with an alternation of quantisation levels in each generation, the reconstruction error is substantially the same for B and I frames and follows a trend to increased error with number of generations but with alternate (odd numbered) generations showing less error than intervening (even numbered) generations.

Comparison of Figure 7 with Figure 8 shows the error is much less than with the conventional MPEG encoding and decoding.

A postulation for the results of Figures 6, 7 and 8 is as follows referring to Figure 9. Figure 9 shows schematically repeated decoding and encoding of an I-frame omitting for simplicity motion vector processing. In Figure 9, I and ε

represent transformed data and I' and ε' represent decoded data. Whilst the following description relates Figure 9 to Figure 1, Figure 9 is in fact a general description of several generations of encoding and decoding.

It will be appreciated that in the embodiments of Figures 4 and 5, the frames which are quantized are I-frames. The predicted frame S^Tp is an I-frame subject, in the encoder, to repeated encoding, decoding and encoding. The original I-frame which is converted to the signal S^Ti which is also an I-frame is subject to repeated encoding and decoding as it passes from Server 1 to Server 2.

At the first stage of coding, corresponding for example to processing of video before storing I-frames in server 1, each pixel of an I frame is subjected to a quantisation process which introduces a quantisation error of up to $\pm\varepsilon/2$ and shown in Figure 9 as ε_1 and ε_2 for the first two generations. The first generation corresponds for example to encoding to produce the intra-frames stored on server 1 and decoding in decoder 4 of Figure 1. The second generation corresponds for example to encoding in encoder 5 and decoding in decoder 6 of Figure 1. Provided the motion prediction values are consistent between generations and the DCT transform and the inverse DCT transform are sufficiently accurate to ensure near perfect reconstruction, then the errors input into the second generation quantiser are the same as those created in the first generation (ε_1). The second quantiser creates a new set of errors ε_2. The output of the quantiser of generation 2 is shown as the original source I added to the separate error components ε_1 and ε_2. The reason for keeping the errors as separate components becomes apparent in the third generation. The third generation corresponds for example to encoder 7 of Figure 1 and subsequent decoding for display. At the end of the third generation, the output signal is almost identical to that provided as the source to the first generation. Therefore, given sufficient accuracy in the transforms and motion vector compensation process, the input to the third generation quantiser will be the transformed frame I of the first generation together with the added errors created by the first and second generations. Now the component $(I'+\varepsilon'_1)$ is a rounded number which needs at least $\pm\varepsilon/2$ additional signal to force a change of state in the quantiser output.

Referring to Figure 6 since the quantiser level is kept constant over the

generations in the first mode of operation, then the error ε_2 created by the second generation quantiser is zero because the quantised levels $I+\varepsilon_1$ fall exactly on the levels of the second quantiser (i.e. $\varepsilon_2=\varepsilon_3=0$). Similarly, any error created by the third generation is cancelled. Thus perfect reconstructions is assured following the first generation loss.

Referring to Figure 7, the quantization values are changed in the second generation, corresponding to change in encoder 5 or quantization to match the bit rate of the channel 3. Thus, the quantiser in the second generation creates a new set of errors ε_2 and the quantised signal is now $I+\varepsilon_1+\varepsilon_2$ and the reconstructed signal $I'+\varepsilon_1'+\varepsilon_2'$ indicates a reduced signal to noise ratio. In the third generation, corresponding to the decoding in the decoder 6 and re-encoding as I in encoder 7, the quantization is set to be the same as the quantization in the first generation. At the input to the third generation, the signal is $I'+\varepsilon_1'+\varepsilon_2'$. If the third generation quantises the signal with the same quantisation level as the first generation then the reconstructed signal out of the third generation might seem to be; $I_1'+\varepsilon_1'+\varepsilon_2'+\varepsilon_3'$ but $\varepsilon_3' = \varepsilon_1'$ and since $I_1'+\varepsilon_1$ is a quantised value, ε_3' is zero. Therefore the output from the third quantiser is $I_1/q_1+\varepsilon_1/q_1+\varepsilon_2/q_1$.

$I_1/q_1+\varepsilon_1/q_1$ is an integer value. In many cases ε_2/q_1 is zero because the error $\varepsilon_2<\varepsilon_1$. Therefore the output signal from the third codec is equal to the output of the first codec plus a small additional error ε_2 when ε_2 exceeds ε_1. The effect is a recovery in S/N ratio at the third generation but with a slight loss over the first generation dependent on the input signal I' and the relationship between the quantisers of the first and second codecs.

In the present invention, the difference data Δ is produced by the subtracting the transformed and quantized data $S^T_{iq} - S^T_{pq}$ which are separately subject to the same quantization. In the conventional MPEG encoder the difference data is produced by subtracting the input and predicted data $S_i - S_p$ and then transforming and quantizing to form $(S_i - S_p)^T_q$.

Forming $S^T_{iq} - S^T_{pq}$ slightly increases the entropy compared to forming $(S_i - S_p)^T_q$.

A further application of the invention is shown in Figures 11 and 12.

Consider by way of example two encoded video bitstreams C and D each having a GOP of two consisting of an I-frame and a B-frame. Such a bit stream is used in the SONY SX (Trademark) video recorder. Assume for example it is desired to edit the bitstreams to the accuracy of one frame by a simple cut at boundary T1 to produce the bitstream E. At least the B-frame 11D is then isolated from its preceding reference frame(s). Furthermore the edited bitstream no longer has an uninterrupted sequence of I-B GOPs, there being adjacent B-frames at the cut T1.

Referring to Figure 12 in accordance with this appication of the invention, a digital bitstream produced by e.g. a camera 121 is encoded using an encoder 122 according to the invention as shown in Figure 4 to produce an encoded bitstream having an I-B GOP of 2. Such a bitstream is for example recorded on tape as in an SX VTR 123 or otherwise transmitted or utilized before being presented for editing in an editor 125. In order to edit two such bitstreams C and D as shown in Figure 11, each bitstream C and D is decoded to baseband using a decoder 124 as shown in Figure 5 and edited at baseband (i.e.unencoded digital video). The edited baseband video is then re-encoded using an encoder 126 as shown in Figure 4 to produce a bitstream F as shown in Figure 11 having the correct sequence of I-B GOPs.

In accordance with one preferred version of this application of the invention, the quantisation levels of the originally encoded frames are maintained constant throughout the decoding and re-encoding. In addition, the frames which are originally encoded as B have motion vectors associated with them. Those motion vectors are also retained (as concealment vectors) at least for those frames which were originally encoded as B and are subsequently re-encoded as B. This preferred version corresponds to the first mode described above.

This preferred version allows substantially loss-less decoding and re-encoding of I and B frames with conversion of I to B and of B to I.

Although Figures 11 and 12 illustrate editing such decoding and re-encoding may occur in other circumstances such as decoding and re-encoding for transmission through a network with subsequent decoding and re-encoding.

Whilst this version of the invention has been described with reference to an I-B GOP of 2, it may be applied to any length of GOP equal to or greater than 2 and

furthermore, the GOP may include P and/or B frames. Any of the I, P and/or B frames may be decoded and re=encoded substantially losslessly as different ones of I, P and/or B.

Also, whilst the bitstream C was decoded to baseband, it could in principle be decoded to I-frames which would also allow frame accurate editing.

Whilst it is preferred to maintain the quantization levels constant as in the first mode the quantisation levels may be changed between generations of encoding as in the second mode described above but with increased degradation over several generations.

Although the invention has been described by way of example with reference to DCT transforming, the principle of the invention is equally applicable to other transforms such as wavelet transform, a sub-band transform, and a pyramid transform all of which are known in the art.

Whilst the encoder of Figure 4 does not comply with the MPEG 2 standard, it produces encoded signals which may be decoded by an MPEG decoder but without the benefit of the reduced reconstruction error in cascaded operations provided by the present invention.

The quantization applied to the frames in the present invention may be linear or non-linear.

Although the invention has been described with reference to frames, it may be applied to fields. Herein and in the claims, the term "pictures" is used generically for fields and frames.

When an inter-encoded picture is produced motion vectors are associated with it. These motion vectors may be conveyed in the bit stream as 'concealment' vectors in the syntax data to avoid recalculation on any subsequent decoding and encoding, and to improve subsequent inter-frame coding accuracy.

CLAIMS

1. A method of encoding video signals in which inter-encoded pictures and intra-encoded pictures are produced from an input video signal comprising a sequence of pictures, wherein each inter-encoded picture is produced by forming the difference between:

a) a transformed and quantised picture of the input signal; and

b) a corresponding separately transformed and quantised predicted picture derived from the input signal;

the quantisations applied to the transformed picture of the input signal and to the transformed predicted picture being the same.

2. A method according to claim 1, wherein the said predicted picture is produced at least in part from a transformed and quantised picture preceding the said first-mentioned transformed and quantised picture of the input signal and which preceding picture is dequantised and inverse transformed.

3. A method according to claim 2 wherein the said predicted picture is a bidirectionally interpolated picture.

4. A method of decoding video signals encoded according to the method of any one of claims 1 to 3, in which each inter-encoded picture is decoded by forming the sum of:

a) the transformed and quantised inter-encoded picture; and

b) a separately transformed and quantised predicted picture;

and dequantising and inverse transforming the sum to produce the decoded picture;

the quantisations applied to the transformed predicted picture and to dequantise the said sum being the same.

5. A method according to claim 4, wherein the predicted picture is produced at least in part from a dequantised and inverse transformed sum preceding the said first-mentioned sum.

6. a method according to claim 5, wherein the predicted picture is a bidirectionally interpolated picture.

7. A method according to any preceding claim, wherein the pictures are transformed using a Discrete Cosine Transform.

8. A method according to any one of preceding claims 1 to 6, wherein the pictures are transformed using a wavelet transform, a sub-band transform or a pyramid transform.

9. A method of processing an input video signal comprising: encoding the input video signal according to the method of any one of claims 1 to 3, 7 or 8, to produce an encoded signal; applying the encoded signal to a transmission channel, storage device or utilization means and receiving the encoded signal therefrom; and decoding the said received encoded signal according to the method of any one of claims 4 to 6, 7 or 8 to produce a decoded picture signal.

10. A method according to claim 9, wherein the said input video signal is produced by dequantisation and inverse transformation of a transformed and quantised intra-encoded signal.

11. A method according to claim 9 or 10, wherein the said decoded picture signal is re-encoded as a transformed and quantised intra-encoded picture signal.

12. A method according to claim 10 or 11, wherein the quantisation levels of the encoded signals which are decoded and encoded are maintained constant.

13. A method of processing an input video signal comprising :

encoding the input video signal according to the method of any one of claims 1 to 3, 7 or 8 to produce an encoded signal;

optionally applying the encoded signal to a transmission channel, storage device, or utilization means and recovering the signal therefrom;

decoding the encoded signal according to the method of any one of claims 4 to 6 7 or 8, to produce the decoded signal;

processing the decoded signal; and

re-encoding the processed decoded signal according to the method of anyone of claims 1 to 3, 7 or 8.

14. A method according to claim 13, wherein the said processing of the decoded signal comprises editing.

15. A method according to claim 13 or 14 wherein the quantisation levels of theeoded signals which are decoded and re-encoded are maintained constant.

16. A method of processing video signals in which at least some intra-encoded pictures are converted to inter-encoded pictures and then back to intra-encoded pictures, the quantization levels of the pictures which are converted and reconverted being maintained constant.

17. A method according to Claim 16 in which the said intra-encoded pictures are converted to inter-encoded pictures using the method of any one of claims 1 to 3 and the converted pictures are reconverted to intra-encoded pictures using the method of any one of claims 4 to 6.

18. A method of encoding video signals substantially as hereinbefore described with reference to Figure 4.

19. A method of decoding video signals substantially as hereinbefore described

with reference to Figure 5.

20. A method of processing video signals substantially as hereinbefore described with reference to Figure 1 optionally together with Figure 4 and/or 5 and optionally with Figure 6 or 7 of the accompanying drawings.

21. A method of processing video signals substantially as hereinbefore described with reference to Figures 11 and 12 of the accompanying drawings.

22. Apparatus for encoding video signals, in which apparatus inter-encoded pictures are produced from an input signal comprising a sequence of pictures, the apparatus including,

means for transforming and quantising the pictures of the input signal,

means for deriving separately transformed and quantised predicted pictures from the input signal, and

means for forming inter-encoded pictures by forming the difference between the transformed and quantised predicted pictures and the transformed and quantised pictures of the input signal,

the quantisations applied to each said transformed picture of the input signal and to the corresponding transformed predicted picture being the same.

23. Apparatus according to claim 22, comprising,

a transform unit for transforming the pictures of the input signal,

a quantiser for quantising the transformed pictures,

means for dequantising the quantised and transformed input pictures,

means for inverse transforming the dequantised pictures,

means for forming predicted pictures from the inverse transformed and dequantised pictures,

means for transforming the predicted pictures,

means for quantising the transformed predicted pictures,

a subtracter for forming the difference between the transformed and quantised

input pictures and the transformed and quantised predicted pictures, and

an entropy encoder for encoding the said differences.

24. Apparatus according to claim 23, further comprising an adder for adding the transformed and quantised predicted pictures to the said differences, to reform the input pictures, the dequantising means dequantising the reformed input pictures.

25. Apparatus according to claim 23 or 24, wherein the predicted picture forming means includes a motion compensator.

26. Apparatus according to anyone of claims 22 to 25, wherein the transforming means performs a Discrete Cosine Transform.

27. Apparatus according to anyone of claims 22 to 25, wherein the transforming means performs a wavelet transform, a sub-band transform or a pyramid transform.

28. Apparatus for decoding video signals encoded by the apparatus of any one of claims 22 to 27, the apparatus including,

means for forming the sums of transformed and quantised inter-encoded pictures received by the apparatus and,

corresponding separately transformed and quantised predicted pictures,

means for dequantising and inverse transforming the sums to produce decoded pictures, and

means for forming quantised and transformed predicted pictures for application to the sum forming means,

the quantisations applied to each predicted pictures and to the corresponding sum being the same.

29. Apparatus according to claim 28, wherein the means for forming the predicted pictures includes a motion compensator coupled to receive the said decoded pictures.

30. A system for processing input video signals comprising,

an encoding apparatus according to any one of claims 22 to 27, to produce encoded pictures,

a transmission channel, storage device, or utilization means coupled to receive the encoded pictures, and

a decoding apparatus according to claim 28 or 29, coupled to receive the encoded pictures from the transmission channel, storage device or utilization means for decoding the encoded pictures to produce decoded picture signals.

31. A system according to claim 30, further comprising means for decoding by dequantisation and inverse transformation, intra-encoded pictures to produce the said input video signals.

32. A system according to claim 31 or 32, further comprising means for re-encoding, by transformation and quantisation, the said decoded picture signals as intra-encoded pictures.

33. A system according to claim 31 or 32, wherein the quantisations of the inter-encoded pictures which are decoded and re-encoded as intra-encoded pictures are maintained constant.

34. Apparatus for processing an input video signal comprising :

means for encoding the input video signal according to any one of claims 22 to 27 to produce an encoded signal;

optionally a transmission channel, storage device, or utilization means to which the encoded signal is applied and recovered therefrom;

means for decoding the encoded signal according to the apparatus of claim 28 or 29 to produce the decoded signal;

means for processing the decoded signal; and

means for re-encoding the processed decoded signal according to anyone of claims 22 to 27.

35. Apparatus according to claim 34, wherein the said means for processing the decoded signal comprises editing means.

36. A method according to claim 34 or 35 wherein the quantisation levels of the encoded signals which are decoded and re-encoded are maintained constant.

37. A system for processing video signals comprising at least intra-encoded pictures, the system comprising,

means for converting at least some of the intra-encoded pictures to inter-encoded pictures, and

means for re-converting the said inter-encoded pictures back to intra-encoded pictures,

the quantisation levels of the intra-encoded pictures which are converted to inter-encoded pictures and back to intra-encoded pictures being maintained constant.

38. A system according to claim 37, where the converting means includes encoding apparatus according to any one of claims 22 ~~18~~ to 27 ~~22~~, and the re-converting means including decoding apparatus according to claim 28 ~~23~~ or 29 ~~24~~.

39. An apparatus of encoding video signals substantially as hereinbefore described with reference to Figure 4.

40. An apparatus of decoding video signals substantially as hereinbefore described with reference to Figure 5.

41. A system for processing video signals substantially as hereinbefore described with reference to Figure 1 optionally together with Figure 4 and/or 5 and optionally with Figure 6 or 7 of the accompanying drawings.

The Patent Office

21

Application No: GB 9715916.4
Claims searched: 1,22

Examiner: Joe McCann
Date of search: 21 October 1997

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.O): H4F(FRG)

Int Cl (Ed.6): H04N(5/926,7/46,7/50)

Other: Online: WPI,INSPEC

Documents considered to be relevant:

Category	Identity of document and relevant passage		Relevant to claims
X	EP 0540358A2	(VICTOR) - see abstract and figure 3	1,22
X	EP 0535960A2	(TOSHIBA) - see abstract	1,22
X	US 5440344	(ASAMURA ET AL) - see abstract and figure 4	1,22

X Document indicating lack of novelty or inventive step
Y Document indicating lack of inventive step if combined with one or more other documents of same category.
& Member of the same patent family
A Document indicating technological background and/or state of the art.
P Document published on or after the declared priority date but before the filing date of this invention.
E Patent document published on or after, but with priority date earlier than, the filing date of this application.